(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 27 December 2001 (27.12.2001)

PCT

(10) International Publication Number WO 01/98333 A2

(51) International Patent Classification⁷: C07K 14/005

(21) International Application Number: PCT/GB01/02817

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

0015308.0 22 June 2000 (22.06.2000) GB 0024544.9 6 October 2000 (06.10.2000) GB

(71) Applicant (for all designated States except US); CELL-TECH PHARMACEUTICALS LIMITED [GB/GB]; 208 Bath Road Slough, Berkshire SL1 3WE (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PAGE, Mark [GB/GB]; 34 The Swallows, Welwyn Garden City, Hertfordshire AL1 1BY (GB). LI, Jing-Li [GB/GB]; 166 Ravenscroft Road, Beckenham, Kent BR3 4TW (GB). PUMPENS, Paul [LV/LV]; Biomedical Research and Study Centre, University of Lativa, Ratsupites Str. 1, LV-1067 (LV).

(74) Agents: CAMPBELL, Patrick, John, Henry et al.; J. A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5JJ (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

333 A

(54) Title: MODIFICATION OF HEPATITIS B CORE ANTIGEN

(57) Abstract: A protein is provided comprising hepatitis B core antigen (HBcAg) wherein one or more of the four arginine repeats has been deleted, said protein comprising the C-terminal cysteine of HBcAg. The deleted region may be replaced by an epitope from a protein other than HBcAg, in which case the HBcAg acts as a carrier to present the epitope to the immune system. The chimeric protein is useful in prophylactic and therapeutic vaccination of a host, for example against hepatitis B virus.

MODIFICATION OF HEPATITIS B CORE ANTIGEN

The invention relates to modified forms of the core antigen of hepatitis B virus (HBV) and to prophylactic and therapeutic vaccines containing the modified antigen.

5

Background to the invention

HBV remains a major healthcare problem throughout both the developed and developing world. Infection with the virus can result in an acute or chronic disease which in a proportion of cases may lead to hepatocellular carcinoma and death. The virus is double shelled, and its DNA is protected inside a protein structure called the core antigen (HBcAg). The core is surrounded by the envelope protein known as the surface or S antigen (HBsAg).

15 HBcAg is an unusual antigen which can be used as a delivery vehicle for specific peptides to the immune system. The antigen has been used to present T-helper, B and cytotoxic lymphocyte (CTL) epitopes from a variety of viral and bacterial pathogens, including epitopes from the surface antigen of HBV, envelope proteins from hepatitis A and antigens from hepatitis C virus. For a review see Ulrich et al (1998) Advances in Virus Research
20 50 141-182.

HBcAg is an excellent vehicle for the presentation of epitopes due to the molecular structure of the protein, which self-assembles into particles. Each particle is generated from either 180 or 240 copies of a monomeric polypeptide. The polypeptide has 183 or 185 amino acids (aa) depending on the subtype of HBV. The monomer, on reaching an appropriate concentration inside the host cell, forms a particle of approximately 27 nm in diameter. Structural studies have shown that amino acids within the region from residues 68 to 90 form a spiked structure on the surface of the particle which is known as the e1 loop. Two monomers joined by disulphide bonds link to form a dimer spike, the most exposed amino acid being at position 80 (at the centre of the e1 loop).

EP-A-421635 (The Wellcome Foundation Limited) describes modification of the HBV core gene to allow insertion of foreign epitopes into the e1 loop without altering the potential of the protein to from particles. Insertion at this site allows maximum exposure of the inserted epitope on the tip of each spike created by dimers of the protein. As there are approximately 180 (or 240) copies of each monomer per particle, each particle is able to present 180 (or 240) copies of the epitope of interest.

Thus, HBcAg can be used to generate hybrid particles to be used as prophylactic and therapeutic vaccines against infectious diseases. However, initial work has identified a 10 high nucleic acid impurity profile due to the inherent nature of the core protein to bind nucleic acid. The binding of nucleic acid is known to be associated with four arginine repeats found at the C-terminus of the protein. Removal of these repeats using genetic tools has been shown to be feasible and results in the production of particles which do not encapsidate nucleic acid. However, removal of this region appears to reduce the inherent 15 stability of the particle structure.

Summary of the invention

In order to maintain particle stability, whilst overcoming the problem of nucleic acid

20 impurity, the inventors have devised an alternative and novel strategy. The strategy
involves generating a clone in which one or more of the arginine repeats of HBcAg is
removed but in which the C-terminal cysteine is retained. The removal of the arginine
repeats reduces binding of nucleic acid, whilst retention of the C-terminal cysteine allows
the formation of a disulphide bond which in the native structure is important for the

25 formation of a stable particle. The deleted repeat(s) may be replaced with sequences
encoding T-helper, B or CTL epitopes from bacterial or viral pathogens, parasites,
allergens or cancer associated antigens. This is made possible by insertion of a suitable
cloning site in place of the deleted region.

30 Thus, the invention provides a protein comprising HBcAg wherein one or more of the four arginine repeats is absent and a C-terminal cysteine residue is present. An epitope from a

protein other than HBcAg may be present in place of the absent arginine repeat(s). The protein may be incorporated into a pharmaceutical composition for prophylactic or therapeutic vaccination, for example against HBV.

The protein of the invention may comprise a second epitope from a protein other than HBcAg, and the second epitope may be in the el loop of HBcAg. By placing a T-helper epitope in the C-terminus and a B-cell epitope in the el loop, it is possible to enhance the response to the B-cell epitope through intrastructural T-cell help. In addition, the strategy can be used to double the number of a particular epitope on each particle, by cloning the same sequence into both the el loop and the C-terminal region.

Brief description of the drawings

Figure 1: Amino acid sequence of hepatitis B core using the single letter code. The Cterminal sequence (aa135-185) is highlighted to detail the deletion strategy. The 4
arginine (R) repeats are emboldened and underlined for emphasis. Three or four arginine
repeat regions are underlined from aa154-178 or aa146-178 respectively. Deletion of the
underlined regions with insertion of the SpeI restriction site generates constructs encoded
by plasmids pTCR₁₅₄ and pTCR₁₄₆ respectively. pTCR₁₅₄ retains the N-terminal arginine
repeat, and pTCR₁₄₆ has all 4 arginine repeats deleted.

Figure 2: DNA sequence coding for HBcAg and location and orientation of oligonucleotide primers used for PCR. The position of the *Spe*I restriction site is given for oligos MGR371, MGR369 and MGR370 (see Table 1).

25

Figure 3: DNA and amino acid sequences of pre-S2 and S epitopes inserted into core. Figure 3A shows the sequence of aa20-55 of the pre-S2 region of the HBV ayw subtype. Figure 3B shows the sequence of aa110-147 of the S antigen of the adw subtype. Figure 3C shows the sequence of aa110-157 of the S antigen of the adw subtype.

30

Figure 4: Agarose gel electroporesis of inverse PCR fragments. Lanes 1, 2, 3 and 4 =

fragments for pTCR₁₄₆, pTCR₁₅₄, pTCSR₁₄₆ and pTCSR₁₅₄ respectively. Lane 5 = size markers. All fragments are of about 5kb as expected.

Figure 5: Immunoblot analysis of expression of core protein in lysates of *E.coli* bacteria transformed with 3' replacement plasmid constructs. All samples express an anti-core antibody reactive protein of various relative molecular weights depending on presence or absence of replacement sequences and size of replacement. Sample order:

Lane $1 = pTCR_{146} E.coli HB101$

Lane $2 = pTCR_{146}/S110-157 E.coli HB101$

10 Lane $3 = pTCR_{146}/S2-2$ E.coli HB101

Lane $4 = pTCR_{154} E.coli HB101$

Lane $5 = pTCR_{154}/S110-147 E.coli HB101$

Lane $6 = pTCR_{154}/S110-157$ E.coli HB101

Lane $7 = pTCR_{154}/S2-2$ *E.coli* HB101

15 Lane $8 = pTCSR_{146} E.coli HB101$

Lane $9 = pTCSR_{146}/S110-157 E.coli HB101$

Lane $10 = pTCSR_{146}/S2-2$ *E.coli* HB101.

Figure 6: Immunoblot analysis of expression of S sequence in lysates of bacteria 20 transformed with 3' replacement plasmid constructs. Constructs incorporating the S sequences (lanes 2, 4, 5 and 7) are anti-S antibody reactive. Sample order:

Lane $1 = pTCR_{146} E.coli HB101$

Lane $2 = pTCR_{146}/S110-157 E.coli HB101$

Lane $3 = pTCR_{154} E.coli HB101$

25 Lane $4 = pTCR_{154}/S110-147$ E.coli HB101

Lane $5 = pTCR_{154}/S110-157 E.coli HB101$

Lane $6 = pTCSR_{146} E.coli HB101$

Lane $7 = pTCSR_{146}/S110-157 E.coli HB101$

Lane 8 = Pre-stain marker (Novex).

30

Figure 7: Immunoblot analysis of expression of pre-S2 sequence in lysates of bacteria

transformed with 3' replacement plasmid constructs. Constructs incorporating the pre-S2 sequences (lanes 2, 4 and 6) are pre-S2 antibody reactive. Sample order:

Lane $1 = pTCR_{146} E.coli HB101$

Lane $2 = pTCR_{146}/S2-2$ *E.coli* HB101

5 Lane $3 = pTCR_{154} E.coli HB101$

Lane $4 = pTCR_{154}/S2-2$ E.coli HB101

Lane $5 = pTCSR_{146} E.coli HB101$

Lane $6 = pTCSR_{146}/S2-2$ E.coli HB101

Lane 7 = Pre-stain marker (Novex).

10

Figure 8: shows averaged anti-HBc responses in mice immunised with various constructs described in the Examples. The titers were calculated as the negative logarithms of the EC50 (effective concentration, 50%) serum dilution on the basis of sigmoidal doseresponse curves.

15

Detailed description of the invention

The modifications to the HBcAg sequence

- 20 As mentioned above, HBcAg is a protein of 183 or 185 amino acids depending on the subtype of HBV. The extra two amino acids in the 185 form of the protein are located between the first and the second arginine repeats. The sequence of a 185 amino acid form of the protein with a pre-sequence is shown in Figure 1. In Figure 1, the mature HBcAg sequence runs from the Met residue at position 25 to the Cys residue at the extreme C-
- 25 terminus, with the sequence from residues 1 to 24 being the pre-sequence. The four arginine repeats are located at the following positions:

		Position in mature 183 aa sequence	Position in mature 185 aa sequence (see Figure 1)
	first repeat	150-152	150-152
	second repeat	157-159	159-161
	third repeat	164-167	166-169
5	fourth repeat	172-175	174-177

One or more of the arginine repeats is deleted in the protein of the invention. Thus, it is possible to delete one, two, three or all four of the repeats and to delete the first repeat, the second repeat, the third repeat and/or the fourth repeat. Any combination of the four repeats may be deleted. The first repeat is primarily responsible for RNA binding and the second, third and fourth repeats are primarily responsible for DNA binding, and in a preferred embodiment the first repeat is retained and the second to fourth repeats are deleted in order to specifically reduce DNA binding.

A sequence lying between residues 145 and 182 of HBcAg is generally absent in the proteins of the invention, and preferably a sequence lying between residues 150 and 177 is absent. The deleted sequence may comprise the whole of the sequence from residue 145 to residue 182 (or from residue 150 to residue 177) or may comprise only a part of the sequence between those residues. Equally, the deleted sequence may extend on either side of those residues. As used herein, expressions such as "a sequence lying between residues x and y is absent" mean that the sequence which is absent may include residues x and y. Removal of sequence upstream of residue 145 may interfere with the particle-forming ability of the protein and is therefore generally not recommended. In 185 aa forms of HBcAg the deleted sequence may end at residue 184, and in 183 aa forms it may end at residue 182.

The C-terminal cysteine residue in the protein of the invention is typically the natural residue from the C-terminus of HBcAg and is typically preceded by the sequence immediately upstream of the residue in HBcAg. The preceding HBcAg sequence may

comprise from 1 to 7 residues, i.e. 1, 2, 3, 4, 5, 6 or 7 residues. Thus, the C-terminus of the protein of the invention may have the sequence Gln Cys, Ser Gln Cys, Glu Ser Gln Cys, Glu Ser Gln Cys, Arg Glu Ser Gln Cys, Ser Arg Glu Ser Gln Cys, Gln Ser Arg Glu Ser Gln Cys or Se. Gln Ser Arg Glu Ser Gln Cys. However, the Cys residue may not be the one from HBcAg;

- 5 in this case, a protein according to the invention may be constructed by truncating the HBcAg sequence and replacing the truncated sequence with another sequence including a Cys residue and optionally an epitope from a protein other than HBcAg. The Cys residue is typically located at the extreme C-terminal end of the protein of the invention but it may be a number of amino acid residues from the extreme C-terminal end. For example, it may
- 10 be from 1 to 20, from 1 to 10 or from 1 to 5 residues from the C-terminus. In any event, the Cys residue must be able to form a disulphide bond.

The protein of the invention typically comprises the following elements linked in an N-terminal to C-terminal direction:

- 15 (i) an N-terminal part of HBcAg which mediates the formation of particles, for example residues 1 to 144 (or 1 to 146 or 1 to 154), and
 - (ii) a C-terminal part of HBcAg comprising the C-terminal cysteine; wherein at least a part of the sequence of HBcAg from between said N-terminal part and said C-terminal part comprising one or more of the arginine repeats is absent.

20

Where the protein also comprises an epitope from a protein other than HBcAg in place of the absent arginine repeat(s), the protein typically comprises the following elements linked in an N- to C-terminal direction:

- (i) an N-terminal part of HBcAg which mediates the formation of particles, for example 25 residues 1 to 144 (or 1 to 146 or 1 to 154),
 - (ii) an epitope from a protein other than HBcAg, and
- (iii) a C-terminal part of HBcAg comprising the C-terminal cysteine; wherein at least part of the sequence of HBcAg between said N-terminal part and said C-terminal part comprising one or more of the arginine repeats is absent and is replaced by said epitope.

Where the protein comprises an epitope from a protein other than HBcAg in the el loop, the protein typically comprises the following elements linked in an N- to C-terminal direction:

- (i) an N-terminal part of the HBcAg sequence comprising e.g. residues 1 to 67 (or 1 to 74 5 or 1 to 79),
 - (ii) an epitope from a protein other than HBcAg,
 - (iii) a second part of the HBcAg sequence comprising e.g. residues 91 to 144 (or 91 to 146,
 - 91 to 154, 86 to 144, 86 to 146, 86 to 154, 80 to 144, 80 to 146 or 80 to 154); and
 - (iv) a third part of the HBcAg sequence comprising the C-terminal cysteine;
- 10 wherein at least a part of the sequence of HBcAg from between residue 145 (or 147 or 155) and the C-terminal cysteine comprising one or more of the arginine repeats is absent.

Where the protein of the invention comprises both a first epitope from a protein other than HBcAg in place of the absent arginine repeat(s) and a second epitope from a protein other

- 15 than HBcAg in the e1 loop, the protein typically comprises the following elements linked in an N- to C-terminal direction:
 - (i) an N-terminal part of the HBcAg sequence comprising e.g. residues 1 to 67 (or 1 to 74 or 1 to 78);
 - (ii) an epitope from a protein other than HBcAg,
- 20 (iii) a second part of the HBcAg sequence comprising e.g. residues 91 to 144 (or 91 to 146, 91 to 154, 86 to 144, 86 to 146, 86 to 154, 80 to 144, 80 to 146 or 80 to 154);
 - (iv) a further epitope from a protein other than HBcAg, and
 - (v) a third part of the HBcAg sequence comprising the C-terminal cysteine; wherein at least a part of the sequence of HBcAg from between residue 145 (or 147 or
- 25 155) and the C-terminal cysteine comprising one or more of the arginine repeats is absent.

As will be apparent from the above, the inventors specifically contemplate modifying the HBcAg sequence in a number of ways, including deletion of one or more of the arginine repeats, insertion of a heterologous epitope in place of the deleted repeat(s) and insertion

30 of a second heterologous in the el loop. However, further modification of the HBcAg sequence is possible. Such further modification may be by way of substitution, insertion,

deletion or extension. The size of an insertion, deletion or extension may, for example, be from 1 to 200 aa, from 1 to 100 aa or from 1 to 50 aa, from 1 to 20 aa or from 1 to 6 aa in the sequence of HBcAg. Substitutions may involve a number of amino acids up to, for example, 1, 2, 5, 10, 20 or 50 amino acids over the length of the HBcAg sequence. The 5 modified protein generally retains the ability to form particles. Substitutions will generally be conservative and may be made, for example, according to the following Table, in which amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

10	ALIPHATIC	Non-polar	GAP
			ILV
		Polar-uncharged	CSTM
			NQ
		Polar-charged	DE
			KR
	AROMATIC		HFWY

Each part of the HBcAg sequence in the protein of the invention preferably has at least 70% sequence identity to the corresponding sequence of a natural HBcAg protein, such as 15 the protein having the sequence shown in SEQ ID NO: 2. More preferably, the identity is at least 80%, at least 90%, at least 98%, at least 97% or at least 99%. Methods of measuring protein sequence (and nucleic acid sequence) identity are well known in the art. For example, the UWGCG Package provides the BESTFIT programme (Devereux et al (1984) Nucleic Acids Research 12, p.387-395). Similarly, the PILEUP and BLAST 20 algorithms can be used to line up sequences (for example as described in Altschul S. F. (1993) J. Mol. Evol. 36:290-300 and Altschul, S. F. et al (1990) J. Mol. Biol. 215:403-10).

The protein of the invention may self-assemble into particles which may closely resemble the particles formed by native HBcAg. The particles may be from 20 to 40 nm in diameter, but are preferably about 27 nm in diameter (which is the size of native HBcAg

particles). They contain no detectable or reduced amounts of nucleic acid (DNA and RNA) compared to particles of native HBcAg. They may contain from 160 to 260 monomers of the protein of the invention, but preferably they contain approximately 180 or approximately 240 monomers (which are the numbers of monomers in native HBcAg particles).

Determination of the particulate nature of a protein according to the invention may be carried out by size exclusion chromatography and/or electron microscopy. Determination of the DNA content of the particles may be carried out by agarose gel electrophoresis or spectrophotometry. A method adapted from Birnbaum and Nasal (1990, J. Virology 64 3319-3330) may be used. The protein may be digested with Proteinase K and the nucleic acid extracted using a commercial DNA recovery kit (e.g. Qiagen, QIAquickTM PCR Purification Kit). Purified DNA may be visualised using a high sensitivity DNA stain (e.g. Novex, SYBER Green ITM) in a 1.5% agarose gel, following electrophoresis. The DNA product obtained following extraction may be quantified using the optical density (OD) 260nm:280nm ratio according to Sambrook *et al.* (1989, Molecular cloning - A laboratory manual, second edition, published by Cold Spring Harbor Laboratory Press), for example using a Pharmacia Biotech Ultraspec 2000TM.

20 The epitopes

As a general rule, epitopes inserted into the protein of the invention should not prevent the folding of HBcAg or its self-assembly into particles. In addition, for improved immunogenicity, B-cell epitopes should be displayed on the surface of the particle. T-cell epitopes do not need to be displayed on the surface of the particle for optimal presentation.

There are three preferred regions for insertion of the epitopes, namely the C-terminus in place of deleted arginine repeat(s), the el loop and the N-terminus. These three regions all tolerate well insertion of foreign sequences. When an epitope is placed in the el loop of HBcAg, it may be inserted in the sequence of amino acid residues 68 to 90, 69 to 90, 71 to 90, 75 to 85 or 78 to 83. Most preferred is to insert the epitope between residues 79 and 80

or 80 and 81. HBcAg residues from the el loop may be deleted in proteins of the invention, so that the inserted epitope may replace all or part of the sequence of the loop.

A heterologous epitope present in a protein of the invention may be a B-cell epitope or a

5 T-cell epitope. In the case that an epitope is a T-cell epitope it may be a T-helper (Th) cell epitope (either a Th1 or Th2 epitope) or a cytotoxic lymphocyte (CTL) epitope.

The protein of the invention may contain more than one heterologous epitope, for example up to 2, 3, 5 or 8 heterologous epitopes, and in this case each epitope may be present in the same site or at different sites in HBcAg. In a preferred embodiment of the invention, one of the epitopes is a T-helper cell epitope and another is a B-cell or a CTL epitope. The presence of the T-helper cell epitope enhances the immune response against the B-cell or CTL epitope. Where there are two or more heterologous epitopes in the protein of the invention, they may be from the same organism or the same protein. Indeed, the epitopes may be the same; this allows a doubling or further multiplication of the number of the epitope presented on the particles.

The size of the sequence comprising an epitope inserted in the protein of the invention can vary between broad limits, but will generally be from 6 to 120 aa, for example from 6 to 20 80 aa or 6 to 40 aa. The epitope may be conformational or linear.

The choice of epitope depends on the disease that it is wished to vaccinate against.

Typically, the epitope is from a pathogen, such as a virus, a bacterium or a protozoan, but it may also be from a cancer associated antigen or an allergen. Examples of pathogens

25 whose epitopes may be inserted include hepatitis A virus (HAV), HBV, hepatitis C virus (HCV), influenza virus, foot-and-mouth disease virus, poliovirus, herpes simplex virus, rabies virus, feline leukemia virus, human immunodeficiency virus type 1 (HIV1), human immunodeficiency virus type 2 (HIV2), simian immunodeficiency virus (SIV), human rhinovirus, dengue virus, yellow fever virus, human papilloma virus, Plasmodium

30 falciparum (a cause of malaria) and bacteria such as Mycobacteria, Bordetella, Salmonella, Escherichia, Vibrio, Haemophilus, Neisseria, Yersinia and Brucella.

Specifically, the bacterium may be Mycobacterium tuberculosis - the cause of tuberculosis; Bordetella pertussis or Bordetella parapertussis - causes of whooping cougl Salmonella typhimurium - the cause of salmonellosis in several animal species; Salmonel... typhi - the cause of human typhoid; Salmonella enteritidis - a cause of food poisoning in 5 humans; Salmonella choleraesuis - a cause of salmonellosis in pigs; Salmonella dublin - a cause of both a systemic and diarrhoeal disease in cattle, especially of new-born calves; Escherichia coli - a cause of food poisoning in humans; Haemophilus influenzae - a cause of meningitis; Neisseria gonorrhoeae - a cause of gonnorrhoeae; Yersinia enterocolitica - the cause of a spectrum of diseases in humans ranging from gastroenteritis to fatal septicemic disease; Brucella abortus - a cause of abortion and infertility in cattle and a condition known as undulant fever in humans; or Clostridium difficile - a cause of pseudomembranous colitis.

Examples of antigens whose epitopes may be inserted are the pre-S1, pre-S2 and S

15 antigens of HBV; the HAV surface antigens; the HCV surface antigens, core protein and

NS3 protein; the HIV antigens gp120, gp160, gag, pol, Nef, Tat and Ref; the malaria

antigens such as the circumpsorozoite proteins; the influenza antigens HA, NP and NA;

the herpes virus antigens EBV gp340, EBV gp85, HSV gB, HSV gD, HSV gH and HSV

early protein; the human papilloma virus antigens E4, E6 and E7; the cancer antigens

20 carcinoembryonic antigen (CEA), P53, ras and myc; the pertactin antigen from Bordetella

pertussis; and house dust mite allergen.

The invention is particularly suited to prophylactic or therapeutic vaccination against HBV since the carrier protein HBcAg is from HBV, and epitopes from the pre-S1, pre-S2 and S regions of HBV are particularly preferred. A pre-S1, pre-S2 or S insert is typically at least 6 amino acids in length, for example from 6 to 120 aa, 8 to 80 aa or 10 to 40 aa. The insert may include, for example, the residues at pre-S1 positions 1-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99, 100-109 or 110-119 or the residues at pre-S2 positions 120-129, 130-139, 140-149, 150-159, 160-169 or 170-174. Particularly preferred

30 fragments are those corresponding to pre-S1 residues 20-47 and pre-S2 residues 139-174. Pre-S1 residues 21-28 correspond to a human T-cell epitope. Also preferred are fragments

corresponding to S residues 110-147 and 110-157 (counting the first residue of the S sequence as residue 1).

Making the proteins of the invention

5

The proteins of the invention are generally made by recombinant DNA technology. The invention includes a nucleic acid molecule (e.g. DNA or RNA) encoding a protein of the invention, such as an expression vector.

The nucleic acid molecule may encode a protein in which one or more of the arginine repeats has been deleted and replaced with a restriction enzyme site unique to the nucleic acid molecule, such as an XbaI site. The nucleic acid molecule may also contain a unique restriction enzyme site in the sequence encoding the el loop and/or in the N-terminus. The unique restriction enzyme sites allow sequences encoding epitopes to be inserted into the nucleic acid molecule, for example in place of the deleted arginine repeat(s) or in the el loop.

A protein of the invention may be produced by culturing a host cell containing a nucleic molecule encoding the protein under conditions in which the protein is expressed, and 20 recovering the protein. Suitable host cells include bacteria such as *E. coli*, yeast, mammalian cell lines and other eukaryotic cell lines, for example insect Sf9 cells.

The vectors constituting nucleic acid molecules according to the invention may be, for example, plasmid or virus vectors. They may contain an origin of replication, a promoter 25 for the expression of the sequence encoding the protein, a regulator of the promoter such as an enhancer, a transcription stop signal, a translation start signal and/or a translation stop signal. The vectors may also contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a neomycin resistance gene for a mammalian vector. Vectors may be used *in vitro*, for example for the production of RNA or used to transform or transfect a host cell. The vector may also be adapted to be used *in vivo*, for example in a method of gene therapy or DNA vaccination.

Promoters, enhancers and other expression regulation signals may be selected to be compatible with the host cell for which the expression vector is designed. For example, prokaryotic promoters may be used, in particular those such as the *trc* promoter suitable for use in *E. coli* strains (such as *E. coli* HB101). A promoter whose activity is induced in 5 response to a change in the surrounding environment, such as anaerobic conditions, may be used. Preferably an *htrA* or *nirB* promoter may be used. These promoters may be used in particular to express the protein in an attenuated bacterium, for example for use as a vaccine. When expression of the protein of the invention is carried out in mammalian cells, either *in vitro* or *in vivo*, mammalian promoters may be used. Tissue-specific promoters, for example hepatocyte cell-specific promoters, may also be used. Viral promoters may also be used, for example the Moloney murine leukaemia virus long terminal repeat (MMLV LTR), the rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, herpes simplex virus promoters and adenovirus promoters. All these promoters are readily available in the art.

15

A protein according to the invention may be purified using conventional techniques for purifying proteins. The protein may, for example, be provided in purified, pure or isolated form. For use in a vaccine, the protein must generally be provided at a high level of purity, for example at a level at which it constitutes more than 80%, more than 90%, more than 20 95% or more than 98% of the protein in the preparation. However, it may be desirable to mix the protein with other proteins in the final vaccine formulation, for example other proteins comprising pre-S1, pre-S2 or S sequence of HBV. The protein is preferably substantially free from nucleic acid (DNA and RNA).

25 Vaccines

The primary use of the proteins of the invention is as therapeutic or prophylactic vaccines.

The invention includes a pharmaceutical composition (e.g. a vaccine composition)

comprising a protein of the invention, a particle of the invention or a nucleic acid molecule

30 of the invention and a pharmaceutically acceptable carrier or diluent.

The principle behind prophylactic vaccination is to induce an immune response in a host so as to generate an immunological memory in the host. This means that, when the host is exposed to the virulent pathogen, it mounts an effective (protective) immune response, i.e an immune response which inactivates and/or kills the pathogen. The invention could form the basis of a prophylactic vaccine against a range of diseases, such as HBV, HAV, HCV, influenza, foot-and-mouth disease, polio, herpes, rabies, AIDS, dengue fever, yellow fever, malaria, tuberculosis, whooping cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis and gonnorrhoeae. The epitopes in the protein of the invention are chosen so as to be appropriate for the disease against which the vaccine is intended to provide protection.

The principle behind therapeutic vaccination is to stimulate the immune system of the host to alleviate or eradicate a disease or condition. There are a number of diseases and conditions which may be susceptible to therapeutic vaccination, such as chronic viral diseases including chronic HBV and chronic HCV, cancer, and allergies such as asthma, atopy, eczema, rhinitis and food allergies.

Chronic viral diseases arise when the immune system of an infected host fails to eliminate the virus, allowing the virus to persist in the host for a long period of time. The invention 20 may be used to induce the immune system of the chronically infected individual so as to eliminate the virus. For example, it is believed that patients with chronic hepatitis have an inadequate T-cell response, and that stimulation of an appropriate T-cell response can eliminate the virus. Thus, in order to treat viral hepatitis using the invention, T-cell epitopes may be inserted into the protein of the invention, such as T-cell epitopes from the 25 pre-S1 and pre-S2 regions of HBV.

Similarly, in the case of cancer, it is believed that enhancement of the T-cell response to tumour antigens may help the immune system to destroy the tumour. It is believed that allergic diseases are caused at least in part by an unbalanced T-cell response in which an inflammatory Th2 responses dominates over an antagonistic Th1 response, and that

allergies may therefore be treated by enhancing the Th1 response. This can be achieved according to the invention by using a protein which stimulates a Th1 response.

More than one protein according to the invention may be administered to a patient.

5 Furthermore, a protein according to the invention may be used in combination with one or more other compositions. For example, in the treatment of chronic HBV a protein according to the invention may be used in combination with interferon gamma, LamivudineTM, or another immunotherapeutic agent such as HepacareTM (formerly known as HepageneTM). The protein according to the invention and the other composition may be administered simultaneously or sequentially.

Suitable carriers and diluents for inclusion in pharmaceutical compositions of the invention are isotonic saline solutions, for example phosphate-buffered saline. The composition will normally include an adjuvant, such as aluminium hydroxide. The composition may be formulated for parenteral, intramuscular, intravenous, intranasal, subcutaneous or transdermal administration. The composition comprises the protein, particles or nucleic acid in a prophylactically or therapeutically effective amount. Typically, the protein or particles are administered at a dose of from 0.01 to 30 µg/kg body weight, preferably from 0.1 to 10 µg/kg, more preferably from 0.1 to 1 µg/kg body weight. The nucleic acid of the invention may be administered directly as a naked nucleic acid construct using techniques known in the art or using vectors known in the art. The amount of nucleic acid administered is typically in the range of from 1 µg to 10 mg, preferably from 100 µg to 1 mg. The vaccine may be given in a single dose schedule or a multiple dose schedule. The routes of administration and doses given above are intended only as a

25 guide, and the route and dose may ultimately be at the discretion of the physician.

Experimental Section

Experiment 1

5 1. Materials and Methods

New plasmid constructs were generated by inverse PCR so that three or four C-terminal arginine repeat regions were deleted and a *SpeI* restriction site was introduced to allow insertion of replacement sequences coding for B and T cell epitopes (Fig. 1).

10

The plasmid templates for the inverse PCR were ptrc/core and ptrc/core-S1 which encode respectively for non-hybrid hepatitis B core and hybrid hepatitis B core containing amino acids 20-47 of the pre-S1 sequence of hepatitis B surface protein inserted between amino acids 79 and 80 of the immunodominant el loop. Three oligonucleotide primers (Table 1 and Fig. 2) were used for the PCR reaction. These primers introduce a unique SpeI restriction site in the PCR fragments. The primers were also designed to generate new fragments that were truncated at residues 146 or 154 but maintained 7 residues of the C-terminus including the terminal cysteine at position 185 which is thought to be important for maintaining particle stability by formation of disulphide bonds (Fig. 1).

20

1.1 Construction of parental truncated plasmids

Using primers MGR371/370 or MGR369/370 (Table 1 and Figure 2), inverse PCR fragments are generated from plasmid templates of ptrc/core or ptrc/core-S1. This procedure removes 69 nucleotides (encoding for 23 amino acids (aa155-177)) and 93 nucleotides (encoding for 31 amino acids (146-177)) respectively. The PCR fragments sizes were confirmed by analysis on agarose gels and then digested with SpeI restriction endonuclease followed by purification on agarose gels and self-ligation to generate plasmids pTCR₁₄₆, pTCR₁₅₄ and pTCSR₁₄₆ and pTCSR₁₅₄. pTCR plasmids are derived from the ptrc/core-S1 templates. The 146 and 154 numbering denotes the amino acid number at the truncation

point. The four parental truncated plasmids were used to transform *E.coli* HB101 cells and positive colonies were tested by diagnostic PCR using oligonucleotide primers MGR61/MGR168. Core protein expression was confirmed by immunoblotting of bacterial cell lysates using a mouse anti-core antibody.

5

1.2 Subcloning of replacement sequences into truncated parental plasmids

Three sequences have been subcloned into the 3' end of the truncated parental plasmids described in section 1.1. These include sequences encoding for amino acids 110-147 and 110-157 of the small hepatitis B surface protein, and aa20-55 of the S2 region of the middle hepatitis B surface protein (Figure 3).

For insertion of the 110-157 sequence (plus 2 amino acids resulting from the *NheI* restriction site) oligonucleotide primers MR245-247 (Table 1B) were used to generate a 15 PCR fragment of 147 nucleotides using pMBdSRE/17 as template (Figure 3). This plasmid encodes for the small hepatitis B surface protein (*adw* subtype) for expression in mammalian cells using the mouse metallothionine promoter.

For insertion of the 110-147 sequence (plus 2 amino acids from the *NheI* site)
20 oligonucleotide primers MGR247/264 (Table 1B) were used to generate a PCR fragment of 120 nucleotides using pMBdSRE/17 as template (Figure 3).

For insertion of the 20-55 sequence (plus 2 residues from the *Nhe*I site) of pre-S2, oligonucleotide primers MGR243/249 (Table 1B) were used to generate a PCR fragment of 114 nucleotides using pMByS2R/8 as template (Figure 3). This plasmid encodes for the middle hepatitis B surface protein (*ayw* subtype) under control of the metallothionine promoter for mammalian cell expression.

The PCR fragments were digested with *NheI* restriction endonuclease and purified on agarose gels. The purified fragments were then ligated with *SpeI* digested, phosphatase treated parental plasmids (section 1.1). *E.coli* HB101 cells were then transformed with the

resulting plasmids and positive colonies tested by diagnostic PCR using oligonucleotide primers MGR61/168, immunoblotting with antibodies specific for the insert and partial DNA sequencing of the inserts.

5 2. Results

2.1 Confirmation of inverse PCR fragment generation

Inverse PCR fragments for pTCR₁₄₆, pTCR₁₅₄, pTCSR₁₄₆ and pTCSR₁₅₄ were analysed by separation on 1% agarose gels (Figure 4). The PCR fragments were found to be of the appropriate size (approx. 5.2kb) and were confirmed to be correct by diagnostic PCR (not shown). Immunoblot analysis showed that the parental constructs and those containing the inserted sequences expressed the core protein that was reactive to an anti-core antibody (Figure 5). Further, confirmation of protein expression of the inserted sequences was shown by immunoblotting using anti-S (Figure 6) and anti-pre-S2 antibodies (Figure 7).

Table 1. Oligonucleotide primers used for inverse and diagnostic PCR

Table 1A

20	Oligos	5'-3' sequence	ptrc/core
	MGR61	CTGCACTCAGGCAAGCCATT	230bp-249bp
	MGR62	GCCGAGGCAGGTCCCCTAGA	530bp-549bp
	MGR168	GAAAATCTTCTCGGATCCGC	from vector (pKK233.2)
	MGR282	AGAGATCTCCATGGATTCAG	-10bp-10bp
25	MGR280	GTGGCTTTGGGGCCATGGACA	60bp-79bp
	MGR369	AGGACTAGTGCCTCGGCCCCGTCGTCT	520bp-546bp
	MGR370	AGAACTAGTCAATCTAGGGAATCTCAA	598bp-624bp
	MGR371	TCTTCTAACACTAGTAGTTTCCGG	502bp-525bp

Bold denotes Spel sites

Table 1B

	Oligos	5'-3' sequence	gene and loca
	MGR245	CAGCTAGCGCAATTTCCATCCGTA	HBsAg 147aa
	MGR247	GTTTGTGCTAGCATTCCAGGAACA	HBsAg 110aa
5	MGR264	CCATAGGTTGCTAGCGAAAGCCCA	HBsAg 157aa
	MGR243	TTGCTAGCGTTCAGCGCAGGGTCC	Pre-S2 20aa
	MGR249	GTGAGAGCTAGCTATTTCCCTGCT	Pre-S2 55aa

Experiment 2

10

Summary

Full-length and C-terminally truncated hepatitis B core antigen (HBc) derivatives, which carried long foreign amino acid insertions at position 144, were constructed.

15 HBV preS1, preS2, and HIV-1 Gag fragments of 50-100 amino acids in length were used as such insertions, and the appropriate recombinant genes were expressed in E.coli cells. The appropriate chimeric HBc and HBcΔ derivatives were purified and examined antigenically and immunogenically. Subclass analysis of the induced anti-HBc immune response in mice showed that the Ig ratio of IgG1, IgG2a, and IgG2b

20 antibodies was restored from the IgG1>IgG2a≥IgG2b pattern, which is typical for C-terminally truncated HBcΔ derivatives, to IgG2a≥IgG2b≥IgG1, which is typical

for full-length HBc derivatives, after immunisation with C-terminally truncated HBc∆

derivatives which carried long C-terminal additions of 50-100 amino acids in length.

25 Materials and Methods

Bacterial Strains

E.coli strains RR1 (F, hsdS20 (r b, mb), recA+, ara-14, proA2, lacY1, galK2, rpsL20 (Smr), xyl-5, mtl-1, supE44, λ), and K802 (hsdR, gal, met, supE, mcrA, mcrB) were 30 used for selection and expression of chimeric genes, respectively.

Animals

BALB/C (H-2^d) female mice were used approximately 7-10 weeks old, weight 20 mg. New Zealand white strain female rabbits were used for obtaining polyclonal antibodies.

5

Construction of HBc Derivatives

Vectors based on plasmids pHBc3 and pHBc16-15. Vector pHBc3 was constructed by putting the HBc gene under the control of the tandem repeat of E.coli trp promoters. Vector pHBc16-10 15 was constructed by insertion of an oligonucleotide linker carrying Cla I/Eco RV restriction sites into position 144 of the HBc gene.

Construction of chimeric HBc derivatives. The structure of the HBc and HBc∆ derivatives is shown in Table 2. The recombinant genes were constructed by insertion of the appropriate 15 HBV preS1, preS2, and HIV-1 gag fragments into the Cla I site of the pHBc16-15 vector, with or without in-frame junction to the C-terminal part of the HBc gene.

Purification of Chimeric HBc Derivatives

E.coli cells were grown overnight on a rotary shaker at 37°C in 750 ml flasks containing 300 ml of M9 minimal medium supplemented with 1% casamino acids (Difco Laboratories, Sparks, USA) and 0.2% glucose. An optical density OD₅₄₀ of 2-5 was usually reached. Generally, cells were pelleted and lysed by 30 min incubation on ice in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 50 μg/ml PMSF, 2 mg/ml lysozyme and then ultrasonicated 3 times for 15 s at 22 kHz. Lysates were then adjusted to 10 mM MgCl₂, and 25 20 μg/ml DNAase. After low speed centrifugation, proteins were precipitated from the supernatant with ammonium sulfate at 33% saturation for 1-2 h at 4°C. Pellets were resuspended in a standard PBS buffer containing 0.1% Triton X-100TM, and 5 ml of the solutions were loaded on a Sepharose CL4BTM column (2.5 x 85 cm) and eluted with PBS buffer without Triton X-100. The presence of HBc polypeptides in fractions was tested by PAGE. Positive fractions were pooled and concentrated by ammonium sulfate precipitation at 33% saturation for 20 h at 4°C. Pellets were resuspended in PBS, or in Tris-saline buffer,

10 mM Tris-HCl (pH7.5), 150 mM NaCl, to a final concentration of about 5-20 mg/ml, dialyzed overnight against 2000 volumes of the same buffer and stored at -70°C or at -20°C in 50% glycerol.

5 Polyacrylamide Gel Electrophoresis and Western Blotting

For PAGE analysis, bacteria were pelleted, suspended in SDS-gel electrophoresis sample buffer containing 2% SDS and 2% 2-mercaptoethanol and lysed by heating at 100°C for 5 min. The proteins were separated by Laemmli's polyacrylamide gel electrophoresis (PAGE) in a slab gel (150x150x0.75 mm) apparatus with a gradient 12-18% running gel and a 4% stacking gel. Western blotting was performed in general as described by Towbin et al (1979) in Proc. Natl. Acad. Sci. USA 76 4350-4354. Nitrocellulose sheets (0.2 μ, Millipore, Bedford, USA) were incubated with anti-HBc antibodies and anti-preS1 antibody in dilutions of 1:100 to 1:1000 overnight and then with anti-mouse IgG peroxidase conjugate (1:1000) for 1-2 h at room temperature. The reaction was developed with 3,3'-diaminobenzidine. In parallel, gels were silver-stained according to Ohsawa and Ebata (1983) Anal. Biochem. 135 409-415.

Immunisations

Mice (five per group) were immunised at day 0 intraperitoneally with 0.02 mg of chimeric particles in complete Freund's adjuvant (CFA, Difco) followed by two booster immunisations in Freund's incomplete adjuvant (IFA, Difco) given at days 10 (0.01 mg intraperitoneally) and 24 (0.01 mg intraperitoneally and 0.01 mg subcutaneously). Sera obtained on day 32 were analysed by ELISA for reactivity with HBc particles.

25 ELISA

For the ELISA, recombinant HBc particles were coated onto 96-well microtiter plates by airdrying in a chemical hood overnight. Wells were blocked with 0.5% BSA in PBS for 1 h, incubated with serial dilutions of the various antibodies for 1 h at 37°C and processed with the appropriate second antibodies conjugated to horse radish peroxidase (Sigma) according to the protocols of the manufacturers. Plates were washed 5 times between incubations with 0.05% Tween-20TM in PBS, and 5 times with distilled water to remove Tween-20. Optical

absorbances were measured at 492 nm in an automatic Immunoscan MS[™] reader. The titres were calculated as the negative logarithms of the EC50 (effective concentration, 50%) seru dilution on the basis of sigmoidal dose-response curves. GraphPad Prism[®] version 3.6. software was used in the mean titre calculations.

5

Results

Immunogenicity of Recombinant Proteins. To measure the immunogenicity of HBc carrier and inserted preS1, preS2, and Gag sequences, individual mice sera were repeatedly tested by direct ELISA using recombinant HBcAg and synthetic preS1, preS2, and HIV-1 p24 peptides on solid support. Immunisation with chimeric particles induced high levels of anti-HBc and relatively low levels of anti-insertion antibodies (not shown).

Induction of Different Immunoglobulin Subclasses by Chimeric HBc∆-preS1(20-47) Particles

15 In order to average obtained immunisation data and to make them more informative for comparative subclass analysis of induced immunoglobulins, we calculated mean titres for each group of immunised animals as the negative logarithms of the EC50 (effective concentration, 50%) serum dilution on the basis of sigmoidal dose-response curves (GraphPad Prism® version 3.02). These data on the anti-HBc response of immunised mice, which allow direct comparison of averaged titres, are given in Fig. 8.

The data presented in Fig. 8 show that the wild type HBcAg induces anti-HBc response with the immunoglobulin subclass distribution IgG2a≥IgG2b>IgG1, whilst the immune response to the C-terminally truncated HBc∆ structure T31 presents the IgG1>IgG2b≥IgG2a subclass 25 distribution pattern. The full-length HBc derivative 10-62, which carries a 50 aa long preS1 insertion, shows a subclass distribution analogous to that of the full-length HBc vector. Moreover, replacement of the C-terminus of the HBc molecule by a long foreign insertion (50 amino acids of the preS1 sequence) in the HBc derivative 10-140 makes the subclass distribution of the anti-HBc antibodies rather similar to that induced by the full-length HBc 30 structure (Fig. 8). The HBc∆ derivative 48-2 with a 100 aa long insertion of HIV-1 Gag

occupies an intermediate position in this sense between wild type HBcAg and C-terminally truncated HBc Δ T31 structures.

Table 2. Structure of HBc derivatives with C-terminal insertions. Amino acids appearing at the HBc and insertion sequence junctions are shown in lowercase.

Full-length HBc derivatives

Constr uct	Inser tion				Seq	uence			
HBc 10-62	preS	НВС	pres:						нвс
	131- 80	144 P krsi	31 skrsis DPA	40 FRANTANPDI	50 WDFNPNKDTW	60 PDANKVGAGA	70 FGLGFTPPHO	08 . GELLGWSPQ	145 s E
HBc 9-	preS	нвс	preS2						HBc
87	2 1- 54	144 P krsi	1 QAMQWNSTTI	10 HQTLQDPR	20 VRGLYFPAGGS	30 SSSGTVNPVP	40 TTVSPISSIF	50 SRIGDPAL	145 ks E

C-terminally truncated HBc derivatives

Constr uct	Inser tion	Sequence .
HBc∆ 10-140	preS 131- 79	HBC <u>preS1</u> 144 31 40 50 60 70 79 P krsiskrsis DPAFRANTANPDWDFNPNKDTWPDANKVGAGAFGLGFTPPHGGLLGWSP hdigdycc
HBc∆ 9-142	preS 2 1- 55	HBC preS2 144 1 10 20 30 40 50 55 P krsi QAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDPALN gdycc
HBcΔ 48-2	HIV p55 121- 210	144 pl7 p24 l P ns DTGHSSQVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTM 56 78 LNTVGGHQAAMQMLKETINEEAA agmqasla

CLAIMS

1. A protein comprising hepatitis B core antigen (HBcAg) wherein one or more of the four arginine repeats is absent and a C-terminal cysteine residue is present.

5

- 2. A protein according to claim 1 wherein a first epitope from a protein other than HBcAg is present in place of the absent arginine repeat(s).
- 3. A protein according to claim 1 or 2 wherein the first arginine repeat is present and the second to fourth arginine repeats are absent.
 - 4. A protein according to any one of the preceding claims wherein a sequence lying between residues 145 and 182 of HBcAg is absent.
- 15 5. A protein according to any one of the preceding claims wherein a sequence lying between residues 150 and 177 of HBcAg is absent.
 - 6. A protein according to any one of the preceding claims which comprises a second epitope from a protein other than HBcAg, the second epitope being in the el loop.

20

- 7. A protein according to claim 6 wherein the second epitope is a B-cell epitope.
- 8. A protein according to any one of claims 2 to 7 wherein the first epitope is a T-cell epitope.

25

- 9. A protein according to claim 8 wherein the first epitope is a T-helper cell epitope and the second epitope is a B-cell epitope.
- 10. A protein according to claim 6 which comprises said first and second epitopes wherein the epitopes are the same.

11. A protein according to any one of claims 2 to 10 wherein the first and/or the second epitope is from hepatitis B virus (HBV).

- 12. A protein according to claim 11 wherein the first and/or the second epitope is from
 the pre-S1, pre-S2 or S region of HBV.
 - 13. A protein according to claim 1 comprising the following elements linked in an N-terminal to C-terminal direction:
 - (i) an N-terminal part of HBcAg which mediates the formation of particles, and
- (ii) a C-terminal part of HBcAg comprising the C-terminal cysteine; wherein at least a part of the sequence of HBcAg from between said N-terminal part and said C-terminal part comprising one or more of the arginine repeats is absent.
- 15 14. A protein according to claim 1 comprising the following elements linked in an N-to C-terminal direction:
 - (i) an N-terminal part of HBcAg which mediates the formation of particles.
 - (ii) an epitope from a protein other than HBcAg, and
 - (iii) a C-terminal part of HBcAg comprising the C-terminal cysteine;
- wherein at least a part of the sequence of HBcAg between said N-terminal part and said C-terminal part comprising one or more of the arginine repeats is absent and is replaced by said epitope.
- 15. A protein according to claim 1 comprising the following elements linked in an Nto C-terminal direction:
 - (i) an N-terminal part of the HBcAg sequence comprising residues 1 to 67,
 - (ii) an epitope from a protein other than HBcAg,
 - (iii) a second part of the HBcAg sequence comprising residues 91 to 144, and
 - (iv) a third part of the HBcAg sequence comprising the C-terminal cysteine;
- wherein at least a part of the sequence of HBcAg from between residue 145 and the C-terminal cysteine comprising one or more of the arginine repeats is absent.

16. A protein according to claim 1 comprising the following elements linked in an N-to C-terminal direction:

- (i) an N-terminal part of the HBcAg sequence comprising residues 1 to 67;
- (ii) an epitope from a protein other than HBcAg,
- 5 (iii) a second part of the HBcAg sequence comprising residues 91 to 144;
 - (iv) a further epitope from a protein other than HBcAg;
 - (v) a third part of the HBcAg sequence comprising the C-terminal cysteine; wherein at least a part of the sequence of HBcAg from between residue 145 and the C-terminal cysteine comprising one or more of the arginine repeats is absent.

10

15

- 17. A particle comprising multiple copies of a protein as claimed in any one of the preceding claims.
- 18. A nucleic acid molecule encoding a protein as claimed in any one of claims 1 to 16.
 - 19. A nucleic acid molecule according to claim 18 which is an expression vector.
 - 20. A host cell transformed or transfected with a nucleic acid molecule as claimed in claim 18 or 19.

20

21. A process for producing a protein as claimed in any one of claims 1 to 16, which process comprises culturing a host cell containing a nucleic acid molecule which encodes the protein under conditions in which the protein is expressed, and recovering the protein.

25

- 22. A nucleic acid molecule encoding a protein as claimed in claim 1 wherein the sequence encoding one or more of the four arginine repeats of HBcAg is deleted and replaced with a restriction enzyme site unique to the nucleic acid molecule.
- 30 23. A pharmaceutical composition comprising a protein as claimed in any one of claims 1 to 16, a particle as claimed in claim 17 or a nucleic acid molecule as

claimed in claim 18 or 19 and a pharmaceutically acceptable carrier or diluent.

24. A protein according to any one of claims 1 to 16, a particle according to claim 17 a nucleic acid molecule according to claim 18 or 19 for use in a method of prophylactic or therapeutic vaccination of the human or animal body.

25. A protein, particle or nucleic acid molecule according to claim 24 for use in a method of prophylactic or therapeutic vaccination of the human or animal body against HBV.

10

5

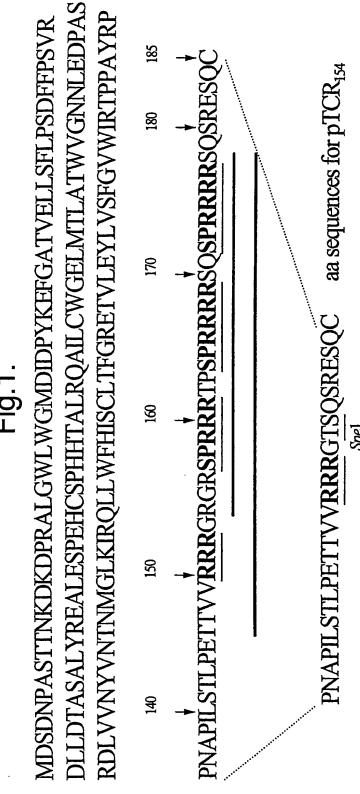
26. Use of a protein according to any one of claims 1 to 16, a particle according to claim 17 or a nucleic acid molecule according to claim 18 or 19 for the manufacture of a medicament for prophylactic or therapeutic vaccination of the human or animal body against HBV.

15

27. A method of vaccination or therapy of a subject, which method comprises administering to the subject a protein as claimed in any one of claims 1 to 16, a particle as claimed in claim 17 or a nucleic acid molecule as claimed in claimed 18 or 19.

20

MIDSDNPASTTNKDKDPRALGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVR Fig.1



aa sequences for pTCR₁₄₆ PNAPIL STLPETSQSRESQC

2/6

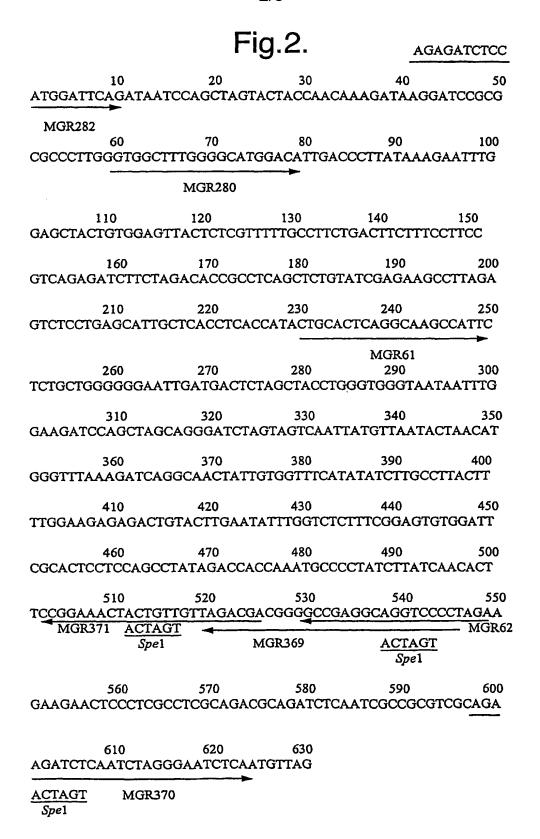


Fig.3A.

CTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTAAACCCTGTTCTGACTACTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCTGCGCTGAAC

Fig.3B.

3/6 ATTCCAGGAACAACAACAACAGGACCATGCAAAACCTGCACGACTCCTGATCTTTCCCTCATGTTGCTGTACAAAACCTTCGGATGGA

AATTGC N C

Ω ပ ပ ပ Σ z 0

AATTGCACCTGTATTCCCATCCCATCGTCTTGGGCT ≽

SUBSTITUTE SHEET (RULE 26)

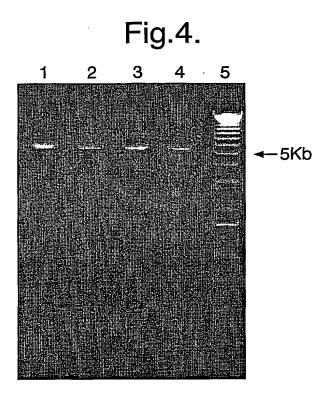


Fig.5.

1 2 3 4 5 6 7 8 9 10

Fig.6.

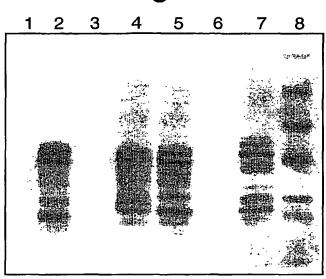
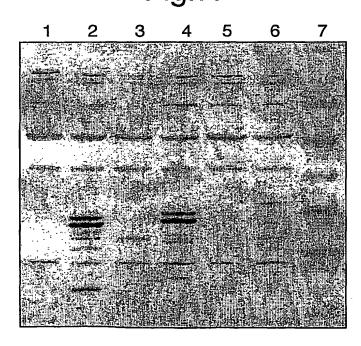
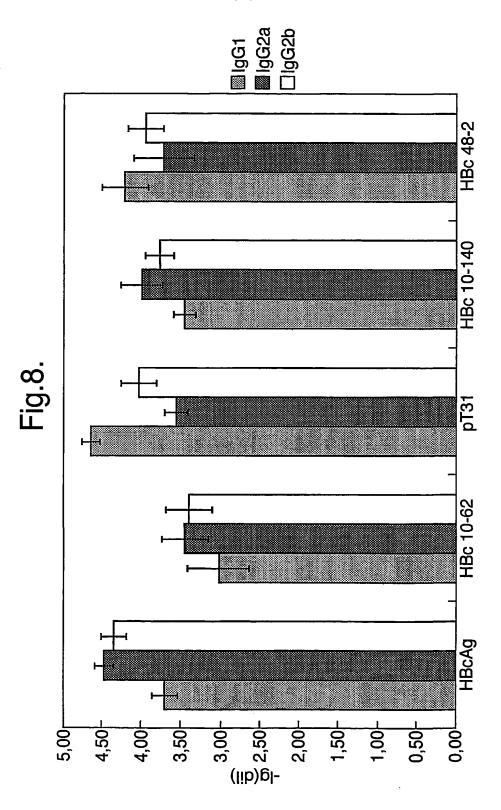


Fig.7.







SEQUENCE LISTING

<11	0> M	EDEV	A EU	ROPE	LTD											
<12)> M	ODIF	ICAT	ION (OF H	EPAT	ITIS	ВС	ORE A	ANTI	GEN					
<13)> P	7845:	1A													
<16)> 2															
<170)> P	aten	tIn '	Ver.	2.1											
<21: <21:	0> 1 1> 6: 2> .DI 3> He	NA	itis	B v	irus											
	L> CI	OS 1)((639))												
<400 ata	_	ctt	ttt	cac	ctc	tac	cta	atc	atc	tct	t.at.	tca	tat	cct	act	48
				His 5												.0
				aag Lys												96
				gaa Glu												144
				ttt Phe												192
				gaa Glu										Pro		240
cat His	act Thr	gca Ala	ctc Leu	agg Arg	caa Gln	gca Ala	att Ile	ctt Leu	tgc Cys	tgg Trp	999 Gly	gaa Glu	cta Leu	atg Met	act Thr	288

				85					90					95		
cta Leu	gct Ala	acc Thr	tgg Trp 100	gtg Val	ggt Gly	gtt Val	aat Asn	ttg Leu 105	gaa Glu	gat Asp	cca Pro	gcg Ala	tct Ser 110	aga Arg	gac Asp	336
cta Leu	gta Val	gtc Val 115	agt Ser	tat Tyr	gtc Val	aac Asn	act Thr 120	aat Asn	atg Met	ggc Gly	cta Leu	aag Lys 125	ttc Phe	agg Arg	caa G1n	384
ctc Leu	ttg Leu 130	tgg Trp	ttt Phe	cac His	att Ile	tct Ser 135	tgt Cys	ctc Leu	act Thr	ttt Phe	gga Gly 140	aga Arg	gaa Glu	aca Thr	gtt Val	432
ata Ile 145	gag Glu	tat Tyr	ttg Leu	gtg Val	tct Ser 150	ttc Phe	gga Gly	gtg Val	tgg Trp	att Ile 155	cgc Arg	act Thr	cct Pro	cca Pro	gct Ala 160	480
tat Tyr	aga Arg	cca Pro	cca Pro	aat Asn 165	gcc Ala	cct Pro	atc Ile	cta Leu	tca Ser 170	Thr	ctt Leu	ccg Pro	gag Glu	act Thr 175	act Thr	528
												act Thr				576
cgc Arg	aga Arg	cga Arg 195	agg Arg	tct Ser	caa G1n	tcg Ser	ccg Pro 200	cgt Arg	cgc Arg	aga Arg	aga Arg	tct Ser 205	caa Gln	tct Ser	cgg Arg	624
G1 u		caa Gln	tgt Cys	tag												639
<212	> 21 > PR	T	tis	B vi	rus											
<400 Met 1		Leu	Phe	His 5	Leu	Cys	Leu	Ile	Ile 10	Ser	Cys	Ser	Cys	Pro 15	Thr	
Val	G1n	Ala	Ser 20	Lys	Leu	Cys	Leu	Gly 25	Trp	Leu	Trp	Gly	Met 30	Asp	Ile	

Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu Ser Phe Leu 35 40 45

- Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp Thr Ala Ser 50 55 60
- Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys Ser Pro His 65 70 75 80
- His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met Thr 85 90 95
- Leu Ala Thr Trp Val Gly Val Asn Leu Glu Asp Pro Ala Ser Arg Asp 100 105 110
- Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys Phe Arg Gln
 115 120 125
- Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg Glu Thr Val 130 135 140
- Ile Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr Pro Pro Ala 145 150 155 160
- Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro Glu Thr Thr 165 170 175
- Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Thr Pro Ser Pro 180 185 190
- Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Ser Gln Ser Arg 195 200 205
- Glu Ser Gln Cys 210

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 27 December 2001 (27.12.2001)

PCT

(10) International Publication Number WO 01/98333 A3

(51) International Patent Classification7: C12N 15/51. C07K 14/02, A61K 39/29, A61P 31/20

(21) International Application Number: PCT/GB01/02817

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

0015308.0 0024544.9

22 June 2000 (22.06.2000) GB 6 October 2000 (06.10.2000)

(71) Applicant (for all designated States except US): CELL-TECH PHARMACEUTICALS LIMITED [GB/GB]; 208 Bath Road Slough, Berkshire SL1 3WE (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PAGE, Mark [GB/GB]; 34 The Swallows, Welwyn Garden City, Hertfordshire AL1 1BY (GB). LI, Jing-Li [GB/GB]; 166 Ravenscroft Road, Beckenham, Kent BR3 4TW (GB). PUMPENS, Paul [LV/LV]; Biomedical Research and Study Centre, University of Lativa, Ratsupites Str. 1, LV-1067 (LV). BORISOVA, Galina [LV/LV]; Biomedical Research and Study Centre, University of Latvia, Ratsupites Str.1, Riga, Latvia 1067 (LV).

(74) Agents: CAMPBELL, Patrick, John, Henry et al.; J. A. Kemp & Co., 14 South Square, Gray's Inn, London WC1R 5JJ (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report

(88) Date of publication of the international search report: 28 March 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MODIFICATION OF HEPATITIS B CORE ANTIGEN

(57) Abstract: A protein is provided comprising hepatitis B core antigen (HBcAg) wherein one or more of the four arginine repeats has been deleted, said protein comprising the C-terminal cysteine of HBcAg. The deleted region may be replaced by an epitope from a protein other than HBcAg, in which case the HBcAg acts as a carrier to present the epitope to the immune system. The chimeric protein is useful in prophylactic and therapeutic vaccination of a host, for example against hepatitis B virus.

Int ational Application No PC I / GB 01/02817

A. CLASSI	FICATION OF SUBJECT MATTER	00 161201/06	······································
170 /	C12N15/51 C07K14/02 A61K39/	29 A61P31/20	
According to	o International Patent Classification (IPC) or to both national classific	ation and IPC	
	SEARCHED	-	
	currentation searched (classification system followed by classificat	ion symbols)	
IPC 7	CO7K		
Documental	ion searched other than minimum documentation to the extent that	such documents are included in the fields si	earched
ŀ			
Electronic d	ala base consulted during the international search (name of data ba	as and where practical peaceh terms used	n
		•	"
F50-TU	ternal, BIOSIS, WPI Data, PAJ, MEDL	INE, CHEM ABS Data	
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT	- W	
Category °	Citation of document, with indication, where appropriate, of the re	levant passages	Relevant to claim No.
χ	HATTON T ET AL: "RNA AND DNA-BI	NDING	1,3-5,
^	ACTIVITIES IN HEPATITIS B VIRUS (13,17-22
	PROTEIN A MODEL FOR THEIR ROLES		
	REPLICATION"		
	JOURNAL OF VIROLOGY,	5041	
	vol. 66, no. 9, 1992, pages 5232- XP001041852	-5241,	
	ISSN: 0022-538X		
	page 5232, left-hand column, para	agraph 2	
	-right-hand column, paragraph 1	29. ap., 2	
	page 5232, right-hand column, page		
	-page 5233, left-hand column, par	ragraph 2	
	page 5234; figure 1		
	page 5234, left-hand column, para -right-hand column, paragraph 1	agraph I	
	page 5235, right-hand column, par	ragraph 3	
	page 5238; figure 7	ug. up., o	
	-	-/	
X Furth	er documents are listed in the continuation of box C.	X Patent family members are listed	in annex.
Special cal	egories of cited documents :	*T* later document published after the inte	rnational filino date
	nt defining the general state of the art which is not	or priority date and not in conflict with cited to understand the principle or the	the application but
	ered to be of particular relevance locument but published on or after the international	invention	, , ,
filing d	ate	 X* document of particular relevance; the c cannot be considered novel or cannot 	be considered to
which i	nt which may throw doubts on priority claim(s) or sciled to establish the publication doe of another	involve an inventive step when the do "Y" document of particular relevance; the c	laimed invention
	or other special reason (as specified) nt referring to an oral disclosure, use, exhibition or	cannot be considered to involve an involve an involve an involve document is combined with one or mo	rentive step when the
other n	neans In published prior to the international filing date but	ments, such combination being obviou in the art.	
	an the priority date claimed	*&* document member of the same patent t	amily
Date of the a	actual completion of the international search	Date of mailing of the international sea	rch report
7	December 2001	21/12/2001	
Name and m	nailing address of the ISA	Authorized officer	
	European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk		
	Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,	Sitch W	

2

In' ational Application No

		PCF/GB 01/02817
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 00 26385 A (POWDERJECT VACCINES INC) 11 May 2000 (2000-05-11) page 2, paragraph 1 page 26 -page 28; example 1 page 28 -page 30; example 2	2
Α	SCHOEDEL F ET AL: "HYBRID HEPATITIS B VIRUS CORE ANTIGEN AS A VACCINE CARRIER MOIETY: I. PRESENTATION OF FOREIGN EPITOPES" JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 44, no. 1/3, January 1996 (1996-01), pages 91-96, XP000891474 ISSN: 0168-1656 page 91 abstract	
Α	EP 0 421 635 A (WELLCOME FOUND) 10 April 1991 (1991-04-10) cited in the application page 2, line 35 - line 46	
A	BORISOVA G P ET AL: "RECOMBINANT CORE PARTICLES OF HEPATITIS B VIRUS EXPOSING FOREIGN ANTIGENIC DETERMINANTS ON THEIR SURFACE" FEBS LETTERS, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 259, no. 1, 1 December 1989 (1989-12-01), pages 121-124, XP002035980 ISSN: 0014-5793 page 122, left-hand column, paragraph 5 -page 124, left-hand column, paragraph 1	
A	HUI ERIC KA-WAI ET AL: "Hepatitis B virus maturation is affected by the incorporation of core proteins having a C-terminal substitution of arginine or lysine stretches." JOURNAL OF GENERAL VIROLOGY, vol. 80, no. 10, October 1999 (1999-10), pages 2661-2671, XP002185072 ISSN: 0022-1317 page 2661 abstract page 2663; figure 1	
	-/	

Int ational Application No
PC I / GB 01/02817

		PCT/GB 0:	1/0281/
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages		Relevant to claim No.
A	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 1992 ZHOU S ET AL: "CYS RESIDUES OF THE HEPATITIS B VIRUS CAPSID PROTEIN ARE NOT ESSENTIAL FOR THE ASSEMBLY OF VIRAL CORE PARTICLES BUT CAN INFLUENCE THEIR STABILITY" Database accession no. PREV199294115400 XP002185073 abstract & JOURNAL OF VIROLOGY, vol. 66, no. 9, 1992, pages 5393-5398, ISSN: 0022-538X		

information on patent family members

PCT/GB 01/02817

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
WO 0026385	A	11-05-2000	AU	2022200 A	22-05-2000
			EP	1119630 A1	01-08-2001
			WO	0026385 A1	11-05-2000
EP 0421635	Α	10-04-1991	AU	626183 B2	23-07-1992
			AU	6269090 A	11-04-1991
			CA	2025598 A1	20-03-1991
			DD	298134 A5	06-02-1992
			DE	69021002 D1	24-08-1995
			DE	69021002 T2	23-11-1995
			DK	421635 T3	27-11-1995
			EP	0421635 A1	10-04-1991
			ES	2075883 T3	16-10-1995
			ΙE	903370 A1	10-04-1991
			JP	3216186 A	24-09-1991
			NZ	235315 A	25-09-1991
			ZA	9007233 A	27-05-1992